

UNIDAD 2:

Teoría de Sistemas

Diccionario de Datos

Introducción: Estructurando la Base del Conocimiento

Luego de comprender la diferencia entre dato e información, y la importancia de transformar datos en información útil, es necesario avanzar hacia una herramienta clave en la organización de sistemas informáticos: el **diccionario de datos**. En el desarrollo de software, los sistemas suelen manejar grandes volúmenes de datos distribuidos en diferentes estructuras, como bases de datos, tablas, y archivos. El diccionario de datos surge como un recurso indispensable para documentar, organizar y estandarizar estos elementos de manera eficiente.

Un **diccionario de datos** es un repositorio centralizado que describe los datos de un sistema: sus nombres, tipos, relaciones, orígenes y restricciones. Es como el "manual" del sistema de datos, que permite a los desarrolladores y equipos de trabajo entender cómo interactúan los elementos y cómo está diseñada la estructura de la información. Por ejemplo, en una aplicación de ventas, el diccionario de datos detallará campos como "Cliente", "Producto", "Precio" y sus respectivas especificaciones.

Este tema no solo tiene relevancia técnica, sino también profesional. Como futuros programadores, utilizarás diccionarios de datos para garantizar la consistencia en el manejo de información, facilitar el mantenimiento de sistemas y colaborar con otros desarrolladores. Además, comprender esta herramienta es fundamental para trabajar con bases de datos relacionales, sistemas de gestión de datos y en proyectos de desarrollo a gran escala.

En esta unidad, exploraremos los componentes principales de un diccionario de datos, su estructura, y cómo crear y gestionar uno en el contexto del desarrollo de software. También veremos ejemplos prácticos que te permitirán aplicar este conocimiento en proyectos reales, asegurando que los sistemas sean claros, coherentes y escalables.

Diccionario de datos

El diccionario de datos es un conjunto de metadatos que contiene las características lógicas y puntuales de los datos que se van a utilizar en el sistema que se programa, incluyendo nombre, descripción, alias, contenido y organización. Los metadatos consisten en información que caracteriza datos. Los metadatos son utilizados para suministrar información sobre datos producidos. En esencia, los metadatos intentan responder a las preguntas quién, que, cuando, donde, por qué y cómo, sobre cada una de las facetas relativas a los datos que se documentan. Es muy importante entender el diccionario de datos ya que es una guía en la que se describe la BD y los objetos que la forman. El diccionario contiene las características lógicas de los sitios donde se almacenan los datos del sistema, incluyendo nombre, descripción, alias, contenido y organización.

Como se sabe existen 6 fases para el desarrollo de los sistemas (Relevamiento, Análisis, Diseño, Desarrollo, Prueba e Implementación), el diccionario de datos se crea en la parte del diseño del sistemas, una vez obtenido los requerimientos podemos determinar qué es lo que se va a hacer, luego de realizar y describir los procesos con los DFD (Diagrama de Flujo de Datos), procedemos a la parte de la administración de los datos, en donde se crea la base de datos, normaliza y relaciona, seguido de ello se realiza el diccionario de datos correspondiente, dependiendo el número de tablas descritas en la Base de Datos.

Existen elementos básicos de estructura de datos que necesitan ser guardados en el diccionario de base de datos. Los datos que conforma un diccionario de datos son los siguientes:

1. **Nombre** - Es el nombre del elemento de datos; debe ser significativo.
2. **Alias** - Cualquier otro nombre que se pueda usar para referirse al elemento de datos. Por ejemplo, el nombre de un elemento de datos puede ser Balance actual, y el alias puede ser Deuda. Solo se incluye el alias si realmente es necesario utilizarlo.
3. **Tipo y tamaño** - se refiere a si el elemento de datos contiene valor numérico, carácter o alfabético. Tamaño se refiere al máximo de caracteres o de dígitos que puede tener el elemento de datos.

4. **Formato de salida o mascara editada** - Indica cómo se presenta el dato al mostrarse en pantalla o al imprimirse en un reporte.
5. **Valor por defecto** - Es el valor que el elemento de datos tiene si no se cambia entrando otro valor.
6. **Cabecera de la columna** - Es el nombre que se presenta en la pantalla o el título del dato en el reporte.
7. **Fuente** - De dónde se origina el valor del elemento de datos. Puede ser una forma, un departamento, otro sistema, etc.
8. **Seguridad** - Identifica los individuos o departamentos que pueden modificar el elemento de datos. Por ejemplo, la línea de crédito puede ser cambiada por el gerente de crédito.
9. **Usuario responsable** - Identifica los usuarios responsables de entrar o cambiar los valores del elemento de datos.
10. **Datos aceptables** - Se especifica el dominio o valores permitidos. Pueden ser valores específicos, una lista de valores, los valores que se encuentren en otro archivo, etc. El valor puede tener reglas de validación; por ejemplo, el salario debe estar entre lo permitido para la posición que el empleado ocupa.
11. **Formula** - Si el valor es el resultado de un cálculo, se muestra la fórmula que se utiliza.
12. **Descripciones o comentarios** - Para proveer información adicional, notas o descripciones.

Notación

Los analistas usan símbolos especiales con la finalidad de no usar demasiada cantidad de texto para la descripción de las relaciones entre datos y mostrar con claridad las relaciones estructurales. En algunos casos se emplean términos diferentes para describir la misma entidad (alias) estos se representan con un signo igual (=) que vincula los datos.

Estructura

Cada parte de la información se debe mantener separada para que pueda ser distinguida de otra. Los elementos de los datos son identificados en dos dimensiones: una

estructura de registro y una estructura de columna. Un registro es un grupo relacionado de campos o filas de tablas. Las columnas son lo que define a cada campo en el registro. Un diccionario de datos tiene más información acerca de las columnas que de las filas, ya que las filas se conforman por datos que no son responsabilidad directa del diccionario de base de datos.

Tipos de datos

Los tipos de datos son parte de la definición de una columna de la base de datos. Estos especifican el tipo de dato que se debe introducir en el campo. Los tipos básicos de datos se dividen en numéricos y alfanuméricos. Cada una de estas dos categorías tiene más divisiones. Los campos numéricos deben ser definidos para especificar si los decimales son válidos para el campo. La definición de los campos alfanuméricos define si el campo acepta un carácter o una cadena de muchos caracteres. Los tamaños de los campos también pueden ser aplicados a datos de tipo numérico. En cuanto a los alfanuméricos, pueden tener formatos impuestos a ellos, como por ejemplo el formato de fecha o una restricción en los posibles valores que acepte el campo.

Restricciones de datos

La diferencia entre restricciones y tipos de datos puede ser difícil de notar. Por ejemplo, una restricción en el ancho de una columna es una restricción de columna. Las restricciones de datos no forman parte de una definición de columna. Las restricciones son reglas que habilitan las relaciones entre los datos y la estructura de la base de datos. Las restricciones se aseguran de que los datos de una tabla no coincidan con registros múltiples de otras o que los registros no estén ligados por falta de información esencial; a esto se le llama integridad de los datos y es ejecutada por las restricciones que evitan que algunas columnas contengan valores vacíos incluso obliga a que haya singularidad en los contenidos de columnas específicas.

Otros elementos

Las bases de datos guardan otros objetos que no son hechos para guardar o definir datos, pero sí para formar una parte integral de las operaciones de la base de datos. Un ejemplo son los índices, los cuales aumentan la velocidad de acceso a los datos y

a las vistas, que son consultas que se representan a sí mismas ante usuario como estructuras de la base de datos. Un esquema es una cuenta de usuario que contiene objetos de datos y que aumenta las medidas de seguridad al restringir las acciones que ciertos usuarios pueden ejecutar. Estos elementos son parte de la estructura operacional de las bases de datos sin la necesidad de ser elementos clave de un diccionario de datos.

Los diccionarios de datos nos sirven:

- Para manejar los detalles en sistemas muy grandes, ya que tienen enormes cantidades de datos, aun en los sistemas más chicos hay gran cantidad de datos.
- Para asignarle un solo significado a cada uno de los elementos y actividades del sistema.
- Para documentar las características del sistema, incluyendo partes o componentes, así como los aspectos que los distinguen. También es necesario saber bajo qué circunstancias se lleva a cabo cada proceso y con qué frecuencia ocurren.
- Para facilitar el análisis de los detalles con la finalidad de evaluar las características y determinar donde efectuar cambios en el sistema. Determina si son necesarias nuevas características o si están en orden los cambios de cualquier tipo.
- Localizar errores y omisiones en el sistema, detectan dificultades, y las presentan en un informe. Aun en los manuales, se revelan errores.

Tipos de diccionarios de datos:

- Diccionario Off-Line: Se ocupa de mantener el diccionario en condiciones. No tiene injerencia en el uso dinámico de los datos.
- Diccionario On-Line: Trabaja junto con el compilador. Impide que el programador defina los datos en el programa y los toma directamente del diccionario. Verifica que los datos nombrados existan en el diccionario. Incorpora al programa, desde el diccionario la definición de los datos. Inconveniente: si uno se

olvida de recompilar, estarán conviviendo datos en la correcta versión actual con otros en una versión superada.

- Diccionario In-Line: No incorpora la definición de datos en el programa, sino que las carga cuando se ejecuta.

Ejemplos

Diccionario de datos (Ejemplo)

Empleado

Campo	Tipo	Long	Descripción	Llave	Validación
CódigoE	Número	10	Código del empleado	PK	
CedulaE	Número	10	Número de cédula del empleado		
Nombre	Cadena	30	Nombre del empleado		
Dirección	Cadena	50	Dirección de residencia		
Teléfono	Número	15	Número telefónico de contacto		
Título	Cadena	30	Título de la profesión		
Instit	Cadena	30	Institución de egreso		
Año	Fecha		Año de egreso		Formato: DD/MM/AAAA
Tipo	Cadena	1	Tipo de empleado dentro de la empresa		P: Planta T: Temporal C: Por contrato
CódigoSU	Cadena	10	Código del empleado supervisor	FK	

Un ejemplo ficticio de conjunto de datos de la siguiente forma:

TPCHE	YFAB	NUMC	AVCNS	PWR
SEAT	2020	4	4,1	140
AUDI	2019	4	6,4	130
BMW	2020	3	7,2	320

En este caso, un usuario que desconozca la base de datos, no sabrá interpretar correctamente el significado de los campos “TPCHE”, “YFAB”, “NUMC” ... Sin embargo, si esta tabla viene asociada con un diccionario de datos, podremos relacionar el meta-dato con el conjunto, tal y como se muestra en la siguiente imagen:

Columna 1:	TPCHE
Título:	marca
Descripción:	Este campo contiene información sobre la marca y modelo de cada vehículo.
Tipo de datos:	string
Columna 2:	YFAB
Título:	año
Descripción:	Este campo contiene información sobre el año de fabricación de cada vehículo.
Tipo de datos:	date
Columna 3:	NUMC
Título:	cilindros
Descripción:	Este campo contiene información sobre el número de cilindros de cada vehículo.
Tipo de datos:	integer
Columna 4:	AVCNS
Título:	consumo
Descripción:	Este campo contiene información sobre el consumo medio de cada vehículo, medido en litros / 100 kms.
Tipo de datos:	decimal
Columna 5:	PWR
Título:	potencia
Descripción:	Este campo contiene información sobre la potencia de cada vehículo, medida en CV.
Tipo de datos:	decimal

¿Qué herramientas se utilizan para crear diccionarios de datos?

Existen muchas herramientas y software especializados para crear y gestionar diccionarios de datos. Algunas de las más populares son:

* DBMS (Sistemas de gestión de bases de datos): La mayoría de los DBMS ofrecen funcionalidades para crear y mantener diccionarios de datos.

* Herramientas de modelado de datos: Permiten crear diagramas entidad-relación y generar automáticamente diccionarios de datos.

* Herramientas de gestión de metadatos: Se especializan en el almacenamiento y gestión de metadatos, incluyendo diccionarios de datos.

¿Qué contiene un diccionario de datos?

Podemos definir los siguientes elementos que como mínimo debería contener un diccionario de datos:

- Nombre del elemento: El nombre único que identifica cada campo, tabla o relación.
- Descripción: Una explicación clara y concisa del significado y propósito del elemento.
- Tipo de dato: El formato de los datos almacenados (texto, numérico, fecha, etc.).
- Longitud: El tamaño máximo permitido para los datos.
- Formato: Las reglas que definen cómo se deben presentar los datos.
- Valores permitidos: Los valores válidos que puede tomar un campo.
- Relaciones: Las conexiones entre los diferentes elementos de la base de datos.

Bibliografía

Kendall, K. E. & E. Kendall, J. (2005). *Análisis y diseño de sistemas*: (6 ed.). Pearson Educación. <https://elibro.net/es/ereader/utnfrm/37855?page=1>